



The government seeks individual input; attendees/participants may provide individual advice only.

Middleware and Grid Interagency Coordination (MAGIC) Meeting Minutes¹
February 3, 2020, 12-2 pm ET

Virtual

Participants

Brian Ince (DNI)	Dhruva Chakravorty (Texas A&M U)
Kevin McMorrow (AI Engineering)	Jill Gemmill (Clemson U)
Dan Fay (Microsoft)	David Martin (ANL)
Ravi Madduri (ANL)	Lavanya Ramakrishnan (LBL)
Jessica Li (Illinois)	Alan Sill (TTU and OGF)
Lisa Arafune (Coalition for Academic Scientific Computation)	Kevin L. Thompson (NSF)
Raj Kettimuthu (ANL)	Michael Corn (UCSD)
Tom Gibbs (Nvidia)	Steve Petruzza (Utah)
Hakizumwami Birali Runesha (U of Chicago)	Paul Love (NCO)
Eric Lancon (BNL)	Keith Beattie (LBL)
Sharon Broude Geva (U of Michigan)	Jack Wells (Nvidia)
Kaushik De (U of Texas Arlington)	Kate Keahey (ANL)
Anne Richeson (Century Link)	Joseph Bryan Lyles (ORNL)
Gil Alterovitz (VA)	Rich Carlson (DOE/SC)
Mallory Hinks (NCO)	Saswata Hier-Majumder (DOE/SC)
	Donald Petravick (Illinois)

Proceedings

This meeting was chaired by Richard Carlson (DOE/SC).

Cloud Speaker Series

Guest Speakers:

- Gil Alterovitz, Director of Artificial Intelligence, US Department of Veterans Affairs
- Kate Keahey, Principle Investigator, Chameleon Project, Argonne National Laboratory

¹ Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program.

Recommendations for Leveraging Cloud Computing Resources for Federally Funded Artificial Intelligence Research and Development Report

- Came out in November – publically available now
- Put together a series of recommendations
 - Recommendation 1 – Federal agencies that support AI R&D should launch and support pilot projects to identify and explore the advantages and challenges associated with the use of commercial clouds in conducting federally-funded AI research.
 - VA Data Commons Pilot
 - Recommendation 2 –Task force to outline an approach to educating and equipping those engaged in AI R&D with the skills needed to use cloud-based platforms and technologies.
 - AI@VA Community
 - Recommendation 3 – Task force to examine best practices in identity management and single sign on strategies within government and across the research community to identify approaches that will enable more effective use of the variety of commercial cloud resources.
 - Recommendation 4 – Task force to establish and publish best practices that use open approaches and promote model/code reuse/interoperability among commercial cloud provider platforms
 - AI to go / AI at point of care – enabling R&D models to be pilot tested clinically with cloud backend.

Q&A

- We’re seeing new models being developed for hybrid cloud approaches where we stay on Prem and then we move part of the data on the cloud. Would you share some information on that?
 - That is something being done in the VA Data Commons. In the past there was no cloud at all – concerns about security. Has been some work in having some items local and some on the cloud
- Federated learning. We’re hearing about the prominent role in the medical community on how data is shared and exchanged and the different approaches to federated learning. Can you give an overview of these aspects?
 - There’s a few cases where federated learning really shines. It will be looked at as part of best practices, but it wasn’t specifically cited because some agencies may not need federated learning.
- One of the big cost issues with cloud is moving data in and out. How does that fit in with your data infrastructure?
 - That is thought to be one of the reasons why we will have it centralized. All the data will be moved into this platform and then it will be updated regularly. But it won’t be floating from one medical center to another. VA is doing cost estimates for if you have your own data set and you want to upload or download the results
- Is there a difference in cost between when you put the data in and when you pull it out?
 - That will depend on the terminal cloud provider.

Kate Keahey, Argonne National Laboratory

Chameleon: An Innovation Platform for Computer Science Research and Education

Chameleon

- Trying to create a testbed that adapts itself to your experimental needs
- Cloud++: Leveraging mainstream cloud technologies
 - Powered by OpenStack with bare metal reconfiguration and “special sauce”
- Open, production testbed for computer science research
 - 5,000+ users, 600+ projects, 100+ institutions, 300+ publications

Coming in Phase 3

- Upgrading Haswell to CascadeLake at TACC and UC
- More accelerators: V100s, AMD + AMD GPU, Xilinx
- Storage hierarchy options: Optane, a mix of enterprise NVMe and SSDs
- Composable Hardware
- Networking: replicating FABRIC Hank design, P4 switches
- IoT devices & CHI@Edge

Projects

- Virtualization or containerization – Yuyu Zhou, University of Pittsburgh
 - Testbed requirements
 - Bare metal reconfiguration, isolation, and serial console access
 - Ability to “save your work”
 - Support for large scale experiments
 - Up-to-date hardware
- Classifying cybersecurity attacks – Jessie Walker & team, University of Arkansas at Pine Bluff
 - Testbed requirements
 - Easy to use OpenStack installation
 - A selection of pre-configured images
 - Access to the same infrastructure for multiple collaborators
- Data science research – Blue Keleher, University of Maryland & Emily Herron, Mercer University
 - Testbed requirements
 - Access to distributed storage in various configurations
 - State of the art GPUs
 - Easy to use appliances and orchestration
- Adaptive bitrate video streaming, Divyashri Bhat, UMass Amherst
 - Testbed requirements
 - Distributed testbed facility
 - BYOC – ability to write an SDN controller specific to the experiment
 - Multiple connections between distributed sites

Sharable Experiments

- Jupyter notebooks + Chameleon experimental containers

- Testbed is a “player” for existing experiments/research (compared to a record player)
- Trove to find and share experiments
- Digital publishing with Zenodo

Final Thoughts

- Chameleon is a shareable research instrument as well as a sharing platform

Q&A

- Can you say a few more words about how you’re going to support some concepts in edge computing in the future? Will the user bring their data? Will the user bring their sensor? Are you going to be supporting autonomous instruments?
 - The idea is for it to be Bring Your Own Device. We create a virtual site and you can come and add your edge devices. Those devices could be out there in the wild, where they could have sensors associated with them. Another use is devices that connect via wireless to the experiments.
- Do you have some kind of list of resources?
 - Yes. I mentioned... (audio cut out a bit) It’s probably going to be specific digital artifacts that target reports or types of use cases.

Discussion on LSN Request for Periodic Briefings

- MAGIC is a public group that lives under the umbrella of the Large Scale Networking (LSN) IWG
- LSN wants to make sure that we all understand, in both directions, what is happening in the community. Mallory or Rich will update LSN on MAGIC activities during their monthly meetings.
- Would also like to offer the opportunity to have some more detailed briefings go up to the LSN level. On some periodic basis, once or twice a years, if there is a MAGIC topic or presentation that would be of interest to the larger LSN community, we could suggest that it get put on the LSN agenda for a future meeting.
- If there are topics that are going on at LSN, MAGIC could request to get a presentation on that topic during a MAGIC meeting.

Roundtable

- Jack Wells (NVIDIA) – GPU technology conference is scheduled for April 12-16. Registration is free. Some sessions don’t require registration. Cloud computing is a topic contained within the agenda.
-

Announcements/Action Items

- Mallory will be looking for help drafting a summary of the Workforce Development Speaker Series. Please let her know if you can help put together a short 4-5 page summary of the various workforce development topics.

Next Meeting

March 3 (12 p.m. ET)